2

The Central Idea: The Hilbert Transform

Proofs in this chapter are presented at the end of the chapter.

Prologue: The Hilbert transform is, without question, the most important operator in analysis. It arises in many different contexts, and all these contexts are intertwined in profound and influential ways. What it all comes down to is that there is only one singular integral in dimension 1, and it is the Hilbert transform. The philosophy is that all significant analytic questions reduce to a singular integral; and in the first dimension there is just one choice.

The most important fact about the Hilbert transform is that it is bounded on L^p for $1 < p < \infty$. It is also bounded on various Sobolev and Lipschitz spaces. And also on H^1_{Re} and the space of functions of bounded mean oscillation (*BMO*). We discuss many of these properties in the present chapter and later on in Chapters 4, 5, and 9. See also [KRA5] and [STE2].

Even though the Hilbert transform is well understood today, it continues to be studied intensely. Boundedness properties of the "maximum Hilbert transform" are equivalent to pointwise convergence results for Fourier series. In higher dimensions, the Hilbert transform is used to construct analytic disks. Analytic disks are important in cosmology and other parts of physics.

From our point of view in the present book, the Hilbert transform is important because it is the inspiration and the role model for higher-dimensional singular integrals. Singular integrals in $\mathbb{R}^N$ are about 55 years old. Singular integrals on the Heisenberg group and other more general settings are much newer. We shall study the former in some detail and provide some pointers to the latter. Chapters 9 and 10 develop integral operators on $\mathbb{H}^n$ in some detail—that is one of the main points of this book.

We take it for granted that the reader is familiar with the most basic ideas of Fourier series. Appendix 1 provides a review or quick reference.

S.G.G. Krantz, *Explorations in Harmonic Analysis*, Applied and Numerical Harmonic Analysis, DOI 10.1007/978-0-8176-4669-1_2,

2.1 The Notion of the Hilbert Transform

Capsule: Our first approach to the Hilbert transform will be by way of complex variable theory. The idea is to seek a means of finding the boundary function of the harmonic conjugate of a given function (which in turn is the Poisson integral of some initial boundary function). This very natural process gives rise to a linear operator that may be identified as the Hilbert transform. Later on we shall see that the Hilbert transform arises rather naturally in the context of partial summation operators for Fourier series. Most any question of convergence of Fourier series may be reduced to an assertion about mapping properties of the Hilbert transform. Thus the Hilbert transform becomes a central player in the theory of Fourier series.

Now we study the Hilbert transform H, which is one of the most important linear operators in analysis. It is essentially the *only* singular integral operator in dimension 1, and it comes up decisively in a variety of contexts. The Hilbert transform is the key player—from a certain point of view—in complex variable theory. And it is the key player in the theory of Fourier series. It also comes up in the Cauchy problem and other aspects of partial differential equations.

Put in slightly more technical terms, the Hilbert transform is important for these reasons (among others):

- It interpolates between the real and imaginary parts of a holomorphic function.
- It is the key to all convergence questions for the partial sums of Fourier series.
- It is a paradigm for all singular integral operators on Euclidean space (we shall treat these in Chapter 3).
- It is (on the real line) uniquely determined by its invariance properties with respect to the groups that act naturally on 1-dimensional Euclidean space.

One can *discover* the Hilbert transform by way of complex analysis. As we know, if f is holomorphic on D and continuous up to ∂D, we can calculate f at a point $z \in D$ from the boundary value of f by the following formula:

$$f(z) = \frac{1}{2\pi i} \int_{\partial D} \frac{f(\zeta)}{\zeta - z} \, d\zeta, \quad z \in D.$$

We call

$$\frac{1}{2\pi i} \cdot \frac{d\zeta}{\zeta - z} \tag{2.1.1}$$

the *Cauchy kernel*.

If we let $\zeta = e^{i\psi}$ and $z = re^{i\theta}$, the expression (2.1.1) can be rewritten as follows:

$$\begin{aligned} \frac{1}{2\pi i} \cdot \frac{d\zeta}{\zeta - z} &= \frac{1}{2\pi} \cdot \frac{-i\overline{\zeta} d\zeta}{\overline{\zeta}(\zeta - z)} \\ &= \frac{1}{2\pi} \cdot \frac{-ie^{-i\psi} \cdot ie^{i\psi} d\psi}{e^{-i\psi}(e^{i\psi} - re^{i\theta})} \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2\pi} \cdot \frac{d\psi}{1 - re^{i(\theta-\psi)}} \\
&= \frac{1}{2\pi} \cdot \frac{1 - re^{-i(\theta-\psi)}}{|1 - re^{i(\theta-\psi)}|^2} d\psi \\
&= \left(\frac{1}{2\pi} \cdot \frac{1 - r\cos(\theta - \psi)}{|1 - re^{i(\theta-\psi)}|^2} d\psi \right) \\
&\quad + i \left(\frac{1}{2\pi} \cdot \frac{r\sin(\theta - \psi)}{|1 - re^{i(\theta-\psi)}|^2} d\psi \right). \qquad (2.1.2)
\end{aligned}
$$

If we subtract $\frac{1}{4\pi} d\psi$ from the real part of the Cauchy kernel, we get

$$
\begin{aligned}
\mathrm{Re}\left(\frac{1}{2\pi i} \cdot \frac{d\zeta}{\zeta - z} \right) - \frac{d\psi}{4\pi} &= \frac{1}{2\pi} \left(\frac{1 - r\cos(\theta - \psi)}{|1 - re^{i(\theta-\psi)}|^2} - \frac{1}{2} \right) d\psi \\
&= \frac{1}{2\pi} \left(\frac{\frac{1}{2} - \frac{1}{2}r^2}{1 - 2r\cos(\theta - \psi) + r^2} \right) d\psi \\
&\equiv \frac{1}{2} P_r(e^{i(\theta-\psi)}) d\psi. \qquad (2.1.3)
\end{aligned}
$$

Note that in the last line we have, in effect, "discovered" the classical (and well-known) Poisson kernel.

This is an important lesson, and one to be remembered as the book develops: The real part of the Cauchy kernel is (up to a small correction) the Poisson kernel. That is, the kernel that reproduces harmonic functions is the real part of the kernel that reproduces holomorphic functions.

In the next section we shall examine the imaginary part of the Cauchy kernel and find the Hilbert transform revealed.

2.2 The Guts of the Hilbert Transform

Now let us take the reasoning that we used above (to discover the Poisson kernel) and turn it around. Suppose that we are given a real-valued function $f \in L^2(\partial D)$. Then we can use the Poisson integral formula to produce a function u on D such that $u = f$ (almost everywhere) on ∂D. We may find a harmonic conjugate of u, say $u^\dagger$, such that $u^\dagger(0) = 0$ and $u + iu^\dagger$ is holomorphic on D. What we hope to do is to produce a boundary function $f^\dagger$ for $u^\dagger$. This will create some symmetry in the picture. For we began with a function f from which we created u; now we are extracting $f^\dagger$ from $u^\dagger$. Our ultimate goal is to study the linear operator $f \mapsto f^\dagger$.

The following diagram illustrates the idea:

$$\begin{array}{ccc} L^2(\partial D) \ni f & \longrightarrow & u \\ & & \big\downarrow \\ f^\dagger & \longleftarrow & u^\dagger \end{array}$$

If we define a function h on D as

$$h(z) \equiv \frac{1}{2\pi i} \int_{\partial D} \frac{f(\zeta)}{\zeta - z}\, d\zeta, \quad z \in D,$$

then obviously h is holomorphic in D. We know from calculations in the last section that the real part of h is (up to adjustment by an additive real constant) the Poisson integral u of f. Therefore $\operatorname{Re} h$ is harmonic in D and $\operatorname{Im} h$ is a harmonic conjugate of $\operatorname{Re} h$. Thus, if h is continuous up to the boundary, then we will be able to say that $u^\dagger = \operatorname{Im} h$ and $f^\dagger(e^{i\theta}) = \lim_{r \to 1^-} u^\dagger(re^{i\theta})$.

So let us look at the imaginary part of the Cauchy kernel in (2.1.2):

$$\frac{r \sin(\theta - \psi)}{2\pi |1 - re^{i(\theta - \psi)}|^2}.$$

If we let $r \to 1^-$, then we obtain

$$\begin{aligned} \frac{\sin(\theta - \psi)}{2\pi |1 - e^{i(\theta-\psi)}|^2} &= \frac{\sin(\theta - \psi)}{2\pi (1 - 2\cos(\theta - \psi) + 1)} \\ &= \frac{\sin(\theta - \psi)}{4\pi (1 - \cos(\theta - \psi))} \\ &= \frac{2 \sin(\frac{\theta - \psi}{2}) \cos(\frac{\theta - \psi}{2})}{4\pi \cdot 2\cos^2(\frac{\theta - \psi}{2})} \\ &= \frac{1}{4\pi} \cot\left(\frac{\theta - \psi}{2}\right). \end{aligned}$$

Hence we obtain the Hilbert transform[1] $H : f \to f^\dagger$ as follows:

$$Hf(e^{i\theta}) = \int_0^{2\pi} f(e^{it}) \cot\left(\frac{\theta - t}{2}\right) dt. \tag{2.1.4}$$

[1] There are subtle convergence issues—both pointwise and operator-theoretic—which we momentarily suppress. Details may be found, for instance, in [KAT].

[We suppress the multiplicative constant here because it is of no interest.] Note that we can express the kernel as

$$\cot\frac{\theta}{2} = \frac{\cos\frac{\theta}{2}}{\sin\frac{\theta}{2}} = \frac{1 - \frac{(\theta/2)^2}{2!} \pm \cdots}{\frac{\theta}{2}\left(1 - \frac{(\theta/2)^2}{3!} \pm \cdots\right)} = \frac{2}{\theta}\left(1 + \mathcal{O}(|\theta|^2)\right) = \frac{2}{\theta} + E(\theta),$$

where $E(\theta) = \mathcal{O}(|\theta|)$ is a bounded continuous function. Therefore, we can rewrite (2.1.4) as

$$\begin{aligned} Hf(e^{i\theta}) &\equiv \int_0^{2\pi} f(e^{it}) \cot\left(\frac{\theta - t}{2}\right) dt \\ &= \int_0^{2\pi} f(e^{it}) \frac{2}{\theta - t} dt + \int_0^{2\pi} f(e^{it}) E(\theta - t)\, dt. \end{aligned}$$

Note that the first integral is singular at θ and the second one is bounded and trivial to estimate—just by applying Schur's lemma (see [SAD] and our Lemma A1.5.5).[2]

In practice, we usually write

$$\cot\left(\frac{\theta - t}{2}\right) \approx \frac{2}{\theta - t},$$

simply ignoring the trivial error term. Both sides of this "equation" are called the kernel of the Hilbert transform. When we study the Hilbert transform, we generally use the kernel on the right; and we omit the 2 in the numerator.

2.3 The Laplace Equation and Singular Integrals on $\mathbb{R}^N$

Let us look at Laplace equation in $\mathbb{R}^N$ for $N > 2$:

$$\Delta u(x) = \left(\sum_{j=1}^{N} \frac{\partial^2}{\partial x_j^2}\right) u(x) = 0.$$

The fundamental solution[3] (see [KRA4]) for the Laplacian is

$$\Gamma(x) = c_N \cdot \frac{1}{|x|^{N-2}}, \qquad N > 2,$$

where c_N is a constant that depends on N.

[2] Schur's lemma, in a very basic form, says that convolution with an L^1 kernel is a bounded operator on L^p. This assertion may be verified with elementary estimates from measure theory—exercise.

[3] It must be noted that this formula is not valid in dimension 2. One might guess this, because when $N = 2$ the formula in fact becomes trivial. The correct form for the fundamental solution in dimension 2 is

$$\Gamma(x) = \frac{1}{2\pi} \log |x|.$$

Details may be found in [KRA4].

Exercise for the Reader: Prove the defining property for the fundamental solution, namely, that $\Delta\Gamma(x) = \delta_0$, where δ_0 is the Dirac mass at 0. (**Hint:** Use Green's theorem, or see [KRA4].)

We may obtain one solution u of $\Delta u = f$ by convolving f with Γ:

$$u = f * \Gamma.$$

For notice that $\Delta u = f * \Delta\Gamma = f * \delta_0 = f$.

In the ensuing discussion we shall consider the integrability of expressions like $|x|^\beta$ near the origin (our subsequent discussion of fractional integrals in Chapter 5 will put this matter into context). We shall ultimately think of this kernel as a fractional power (positive or negative) of the fundamental solution for the Laplacian.

The correct way to assess such a situation is to use polar coordinates:

$$\int_{|x|<1} |x|^\beta \, dx = \int_\Sigma \int_0^1 r^\beta \cdot r^{N-1} \, dr \, d\sigma(\xi).$$

A few comments are in order: The symbol Σ denotes the unit sphere in $\mathbb{R}^N$, and $d\sigma$ is rotationally invariant area measure (see Chapter 9 for a consideration of Hausdorff measure on a general surface) on Σ. The factor r^{N-1} is the Jacobian of the change of variables from rectangular coordinates to spherical coordinates. Of course the integral in the rotational variable ξ is trivially a constant. The integral in r converges precisely when $\beta > -N$. Thus we think of $-N$ as the "critical index" for integrability at the origin.

Now let us consider the following transformation:

$$T : f \longmapsto f * \Gamma.$$

The kernel Γ is singular at the origin to order $-(N-2)$. Studying L^p mapping properties of this transformation is easy because Γ is locally integrable. We can perform estimates with easy techniques such as the generalized Minkowski inequality and Schur's lemma (see [SAD] and our Lemmas A1.5.5, A1.5.8). In fact, the operator T is a special instance of a "fractional integral operator." We shall have more to say about this family of operators as the book develops.

The first derivative of Γ is singular at the origin to order $-(N-1)$ and is therefore also locally integrable:

$$\frac{\partial\Gamma}{\partial x_j} = C \cdot \frac{x_j}{|x|^N}.$$

Again, we may study this "fractional integral" using elementary techniques that measure only the *size* of the kernel.

But if we look at the second derivative of Γ, we find that

$$\frac{\partial^2\Gamma}{\partial x_j \partial x_k} = C_{jk} \frac{x_j x_k}{|x|^{N+2}} \equiv K(x)$$

is singular at the origin of order $-N$ and the integral has a critical singularity at 0. Hence, to analyze the transformation

$$\widetilde{T} : f \longmapsto \int f(t)K(x-t)dt,$$

we use the *Cauchy principal value*, denoted by P.V. and defined as follows:

$$\text{P.V.} \int f(t)K(x-t)dt = \lim_{\epsilon \to 0^+} \int_{|x-t|>\epsilon} f(t)K(x-t)dt.$$

We shall be able to see, in what follows, that $\widetilde{T}$ (defined using the Cauchy principal value) induces a distribution. It will also be bounded on $L^p(\mathbb{R}^N)$, $1 < p < \infty$. The operator $\widetilde{T}$ is unbounded on L^1 and unbounded on L^∞. When specialized down to dimension 1, the kernel for the operator $\widetilde{T}$ takes the form

$$K(t) = \frac{1}{t}.$$

This is of course the kernel of the Hilbert transform. In other words, the Hilbert transform is a special case of these fundamental considerations regarding the solution operator for the Laplacian.

In the next section we return to our consideration of the Hilbert transform as a linear operator on function spaces.

2.4 Boundedness of the Hilbert Transform

The Hilbert transform induces a distribution

$$\phi \longmapsto \int \frac{1}{x-t}\phi(t)dt, \qquad \text{for all } \phi \in C_c^\infty.$$

But why is this true? On the face of it, this mapping makes no sense. The integral is not convergent in the sense of the Lebesgue integral (because the kernel $1/(x-t)$ is not integrable). Some further analysis is required in order to understand the claim that this displayed line defines a distribution.

We understand this distribution by way of the Cauchy principal value:

$$\begin{aligned}
\text{P.V.} \int \frac{1}{x-t}\,\phi(t)dt &= \text{P.V.} \int \frac{1}{t}\,\phi(x-t)dt \\
&= \lim_{\epsilon \to 0^+} \int_{|t|>\epsilon} \frac{1}{t}\,\phi(x-t)dt \\
&= \lim_{\epsilon \to 0^+} \left[\int_{1>|t|>\epsilon} \frac{1}{t}\,\phi(x-t)dt + \int_{|t|>1} \frac{1}{t}\,\phi(x-t)dt \right] \\
&= \lim_{\epsilon \to 0^+} \int_{1>|t|>\epsilon} \frac{1}{t}\,[\phi(x-t) - \phi(x)]dt + \int_{|t|>1} \frac{1}{t}\,\phi(x-t)dt.
\end{aligned}$$

In the first integral we have used the key fact that the kernel is odd, so has mean value 0. Hence we may subtract off a constant (and it integrates to 0). Of course the second integral does *not* depend on ϵ, and it converges by Schwarz's inequality.

Since $\phi(x-t)-\phi(x) = \mathcal{O}(|t|)$, the limit in the first integral exists. That is to say, the integrand is bounded so the integral makes good sense. We may perform the following calculation to make this reasoning more rigorous:

For $\epsilon > 0$ define

$$I_\epsilon = \int_{\epsilon<|t|<1} \frac{1}{t} O(|t|)dt.$$

Now if $0 < \epsilon_1 < \epsilon_2 < \infty$ we have

$$I_{\epsilon_2} - I_{\epsilon_1} = \int_{\epsilon_2<|t|<1} \frac{1}{t} O(|t|)dt - \int_{\epsilon_1<|t|<1} \frac{1}{t} O(|t|)dt = \int_{\epsilon_1<|t|<\epsilon_2} \mathcal{O}(1)dt \to 0$$

as $\epsilon_1, \epsilon_2 \to 0$. This shows that our principal value integral converges.

Let $\mathcal{S}$ denote the standard Schwartz space from distribution theory (for which see [STG1], [KRA5]). If $f \in \mathcal{S}$, we have

$$Hf(x) = \text{P.V.} \int \frac{1}{x-t} f(t)dt$$

and

$$\widehat{Hf} = \left(\frac{1}{t}\right)^{\widehat{}} \cdot \widehat{f}.$$

Since $\frac{1}{t}$ is homogeneous of degree -1, we find that $\left(\frac{1}{t}\right)^{\widehat{}}$ is homogeneous of degree 0 (see Chapter 3 on the Fourier transform and Chapter 4 on multipliers). Therefore it is bounded.

Now

$$\|Hf\|_{L^2} = \|\widehat{Hf}\|_{L^2} = \left\| \left(\frac{1}{t}\right)^{\widehat{}} \cdot \widehat{f} \right\|_{L^2} \leq C\|\widehat{f}\|_{L^2} = c\|f\|_{L^2}.$$

By dint of a tricky argument that we shall detail below, Marcel Riesz (and, in its present form, Salomon Bochner) proved that $H : L^p \to L^p$ when p is a positive even integer. By what is now known as the Riesz–Thorin interpolation theorem (stated below), he then showed that H is bounded on $p > 2$. Then a simple duality argument guarantees that H is also bounded on L^p for $1 < p < 2$.

Prelude: Interpolation theory is now an entire subject unto itself. For many years it was a collection of isolated results known only to a few experts. The seminal paper [CAL] cemented the *complex method* of interpolation (the one used to prove Riesz–Thorin) as an independent entity. In the same year, Lions and Peetre [LIP] inaugurated the real method of interpolation. The book [BERL] gives an overview of the subject of interpolation.

In general the setup is this: One has Banach spaces X_0, X_1 and Y_0, Y_1 and an operator

$$T : X_0 \cap X_1 \to Y_0 \cup Y_1.$$

One hypothesizes that

$$\|Tf\|_{Y_j} \leq C_j \|f\|_{X_j}$$

for $j = 0, 1$. The job then is to identify certain "intermediate spaces" and conclude that T is bounded in norm on those intermediate spaces.

Theorem 2.4.1 (Riesz–Thorin interpolation theorem) *Let $1 \leq p_0 < p_1 \leq \infty$. Let T be a linear operator that is bounded on L^{p_0} and L^{p_1}, i.e.,*

$$\|Tf\|_{L^{p_0}} \leq C_0 \|f\|_{L^{p_0}},$$

$$\|Tf\|_{L^{p_1}} \leq C_1 \|f\|_{L^{p_1}}.$$

Then T is a bounded operator on L^p, $\forall p_0 \leq p \leq p_1$, and

$$\|Tf\|_{L^p} \leq C_0^{\frac{p_1-p}{p_1-p_0}} \cdot C_1^{\frac{p-p_0}{p_1-p_0}} \cdot \|f\|_{L^p}.$$

Now let us relate the Hilbert transform to Fourier series. We begin by returning to the idea of the Hilbert transform as a multiplier operator. Indeed, let $\mathbf{h} = \{h_j\}$, with $h_j = -i \operatorname{sgn} j$; here the convention is that

$$\operatorname{sgn} x = \begin{cases} -1 & \text{if } x < 0, \\ 0 & \text{if } x = 0, \\ 1 & \text{if } x > 0. \end{cases}$$

Then the Hilbert transform H is given by the multiplier h. This means that for $f \in L^1(\mathbb{T})$,

$$Hf = \sum_j h_j \widehat{f}(j) e^{ijt}.$$

[How might we check this assertion? You may calculate both the left-hand side and the right-hand side of this last equation when $f(t) = \cos jt$. The answer will be $\sin jt$ for every j, just as it should be—because $\sin jt$ is the boundary function for the harmonic conjugate of the Poisson integral of $\cos jt$. Likewise when $f(t) = \sin jt$ (then Hf as written here is $\cos jt$). That is enough information—by the Stone–Weierstrass theorem—to yield the result.] In the sequel we shall indicate this relationship by $H = \mathcal{M}_h$.

So defined, the Hilbert transform has the following connection with the partial sum operators:

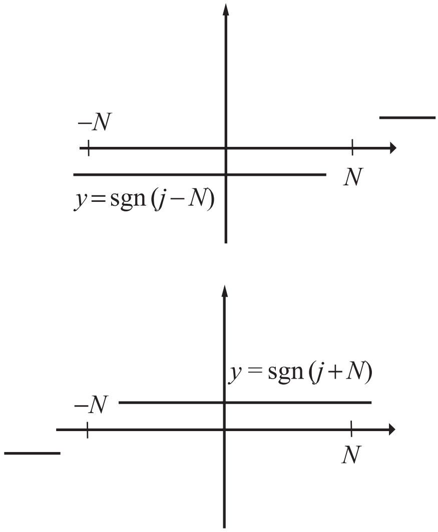

Figure 2.1. Summation operators and the Hilbert transform.

$$\begin{aligned}\chi_{[-N,N]}(j) &= \frac{1}{2}[1+\operatorname{sgn}(j+N)] - \frac{1}{2}[1+\operatorname{sgn}(j-N)] \\ &\quad + \frac{1}{2}[\chi_{\{-N\}}(j) + \chi_{\{N\}}(j)] \\ &= \frac{1}{2}[\operatorname{sgn}(j+N) - \operatorname{sgn}(j-N)] + \frac{1}{2}[\chi_{\{-N\}}(j) + \chi_{\{N\}}(j)].\end{aligned}$$

See Figure 2.1. Therefore, letting $e_k g(t) \equiv e^{ikt} g(t)$ and letting P_j be orthogonal projection onto the space spanned by e^{ijt}, we have

$$\begin{aligned}S_N f(e^{it}) &= \mathcal{M}_{\chi_{[-N,N]}} f(e^{it}) \\ &= i e_{-N} H[e_N f] - i e_N H[e_{-N} f] + \frac{1}{2}[P_{-N} f + P_N f]. \qquad (2.1.5)\end{aligned}$$

To understand this last equality, let us examine a piece of it. We look at the linear operator corresponding to the multiplier

$$m(j) \equiv \operatorname{sgn}(j+N).$$

Let $f(t) \sim \sum_{j=-\infty}^{\infty} \widehat{f}(j) e^{itj}$. Then

$$\begin{aligned}\mathcal{M}_m f(t) &= \sum_j \operatorname{sgn}(j+N) \widehat{f}(j) e^{ijt} \\ &= \sum_j \operatorname{sgn}(j) e^{-iNt} \widehat{f}(j-N) e^{ijt}\end{aligned}$$

$$= ie^{-iNt}\sum_j (-i)\mathrm{sgn}(j)\widehat{f}(j-N)e^{ijt}$$

$$= ie^{-iNt}\sum_j (-i)\mathrm{sgn}(j)(e_N f)\widehat{\ }(j)e^{ijt}$$

$$= ie^{-iNt}H[e_N f](t).$$

This is of course precisely what is asserted in the first half of the right-hand side of (2.1.5).

We know that the Hilbert transform is bounded on L^2 because it is a multiplier operator coming from a bounded sequence. It also turns out to be bounded on L^p for $1 < p < \infty$. [We shall discuss this fact about H below, and eventually prove it.] Similar remarks apply to the projection operators P_j. Taking these boundedness assertions for granted, we now reexamine equation (2.1.5). Multiplication by a complex exponential does not change the size of an L^p function (in technical language, it is an *isometry* of L^p). So (2.1.5) tells us that S_N is a difference of compositions of operators, all of which are bounded on L^p. And the norm is plainly bounded independent of N. In conclusion, if we assume that H is bounded on L^p, $1 < p < \infty$, then Functional Analysis Principle I (see Appendix 1) tells us (since trigonometric polynomials are dense in L^p for $1 \le p < \infty$) that norm convergence holds in L^p for $1 < p < \infty$. We now state this as a theorem:

Prelude: What is remarkable about this next theorem is that it reduces a question of convergence of a sequence of operators to the question of the boundedness of a *single* operator. This illustrates the power of functional analysis—a power that was virtually *discovered* in the context of Fourier analysis. From our modern perspective, the uniform boundedness principle makes this all quite natural.

Theorem 2.4.2 *Fix $1 < p < \infty$ and assume (to be proved below) that the Hilbert transform H is bounded on $L^p(\mathbb{T})$. Let $f \in L^p(\mathbb{T})$. Then $\|S_N f - f\|_{L^p} \to 0$ as $N \to \infty$. Explicitly,*

$$\lim_{N\to\infty}\left[\int_{\mathbb{T}} |S_N f(x) - f(x)|^p\,dx\right]^{1/p} = 0.$$

The converse of this theorem is true as well, and can be proved by even easier arguments. We leave the details to the reader—or see [KAT].

It is useful in the study of the Hilbert transform to be able to express it explicitly as an integral operator. The next lemma is of great utility in this regard.

Prelude: The next lemma is one of the key ideas in Laurent Schwartz's [SCH] distribution theory. It is an intuitively appealing idea that any translation-invariant operator is given by convolution with a kernel, but if one restricts attention to just *functions*, then one will not always be able to find this kernel. Distributions make possible a new, powerful statement.

Lemma 2.4.3 *If the Fourier multiplier $\Lambda = \{\lambda_j\}_{j=-\infty}^{\infty}$ induces a bounded operator $\mathcal{M}_\Lambda$ on L^p, then the operator is given by a convolution kernel $K = K_\Lambda$. In other words,*

$$\mathcal{M}_\Lambda f(x) = f * K(x) = \frac{1}{2\pi}\int_0^{2\pi} f(t)K(x-t)dt.$$

This convolution kernel is specified by the formula

$$K(e^{it}) = \sum_{j=-\infty}^{\infty} \lambda_j e^{it}.$$

[In actuality, the sum that defines this kernel may have to be interpreted using a summability technique, or using distribution theory, or both. In practice we shall always be able to calculate the kernel with our bare hands when we need to do so. So this lemma will play a tacit role in our work.]

If we apply Lemma 2.4.3 directly to the multiplier for the Hilbert transform, we obtain the formal series

$$k(e^{it}) \equiv \sum_{j=-\infty}^{\infty} -i \cdot \operatorname{sgn} j \cdot e^{ijt}.$$

Of course the terms of this series do not tend to zero, so this series does not converge in any conventional sense. Instead we use Abel summation (i.e., summation with factors of $r^{|j|}$, $0 \le r < 1$) to interpret the series: For $0 \le r < 1$ let

$$k_r(e^{it}) = \sum_{j=-\infty}^{\infty} -ir^{|j|} \cdot \operatorname{sgn} j \cdot e^{ijt}.$$

The sum over the positive indices is

$$\begin{aligned}
-i\sum_{j=1}^{\infty} r^j \cdot e^{ijt} &= -i\sum_{j=1}^{\infty}[re^{it}]^j \\
&= -i\left[\frac{1}{1-re^{it}} - 1\right] \\
&= \frac{-ire^{it}}{1-re^{it}}.
\end{aligned}$$

Similarly, the sum over negative indices can be calculated to be equal to

$$\frac{ire^{-it}}{1-re^{-it}}.$$

Adding these two pieces yields that

$$\begin{aligned}
k_r(e^{it}) &= \frac{-ire^{it}}{1-re^{it}} + \frac{ire^{-it}}{1-re^{-it}} \\
&= \frac{-ir[e^{it}-e^{-it}]}{|1-re^{it}|^2} \\
&= \frac{2r\sin t}{|1-re^{it}|^2} \\
&= \frac{2r\sin t}{1+r^2-2r\cos t} \\
&= \frac{2r\cdot 2\cdot \sin\frac{t}{2}\cos\frac{t}{2}}{(1+r^2-2r)+2r(1-\cos^2\frac{t}{2}+\sin^2\frac{t}{2})} \\
&= \frac{4r\sin\frac{t}{2}\cos\frac{t}{2}}{(1+r^2-2r)+2r(2\sin^2\frac{t}{2})}.
\end{aligned}$$

We formally let $r \to 1^-$ to obtain the kernel

$$k(e^{it}) = \frac{\sin\frac{t}{2}\cos\frac{t}{2}}{\sin^2\frac{t}{2}} = \cot\frac{t}{2}. \tag{2.1.6}$$

This is the standard formula for the kernel of the Hilbert transform—just as we derived it by different means in the context of complex analysis. Now we have given a second derivation using Fourier analysis ideas. It should be noted that we have suppressed various subtleties about the validity of Abel summation in this context, as well as issues concerning the fact that the kernel k is not integrable (near the origin, $\cot\frac{t}{2} \approx 2/t$). For the full story, consult [KAT].

Just to repeat, we resolve the nonintegrability problem for the integral kernel k in (2.1.6) using the so-called *Cauchy principal value*, and it will now be defined again. Thus we usually write

$$\text{P.V. } \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)\cot\left(\frac{t}{2}\right)dt,$$

and we interpret this to mean

$$\lim_{\epsilon\to 0^+}\frac{1}{2\pi}\int_{\epsilon<|t|\le\pi} f(x-t)\cot\left(\frac{t}{2}\right)dt. \tag{2.1.7}$$

Observe in (2.1.7) that for $\epsilon > 0$ fixed, $\cot(t/2)$ is actually *bounded* on the domain of integration. Therefore the integral in (2.1.7) makes sense, by Hölder's inequality, as long as $f \in L^p$ for some $1 \le p \le \infty$. The deeper question is whether the limit exists, and whether that limit defines an L^p function.

We will prove the L^p-boundedness of the Hilbert transform, using a method of S. Bochner, below.

The reduction of norm convergence of Fourier series to the study of the Hilbert transform is fundamental to the study of Fourier series. But it also holds great philosophical significance in the modern history of analysis. For it shows that we may reduce the study of the (infinitely many) partial sums of the Fourier series of a function to the study of a *single* integral operator. The device for making this reduction is—rather than study one function at a time—to study an entire space of functions at once. This is what functional analysis is all about.

Many of the basic ideas in functional analysis—including the uniform boundedness principle, the open mapping theorem, and the Hahn–Banach theorem—grew out of questions of Fourier analysis. Even today, Fourier analysis has led to many new ideas in Hilbert and Banach space theory—see [STE2], especially the Cotlar–Knapp–Stein lemma (see Section 9.10).

In the next section we shall examine the Hilbert transform from another point of view.

In the present section, we have taken the validity of Theorem 2.1.2 for granted. The details of this result, and its proof, will be treated as the book develops. Our intention in the next section is to discuss these theorems, and to look at some examples. In the next section we prove the L^p-boundedness of the Hilbert transform.

2.5 L^p Boundedness of the Hilbert Transform

Now we shall prove (at the end of the chapter) that the Hilbert transform is bounded on $L^p(\mathbb{T})$, $1 < p < \infty$. We will present an argument due to S. Bochner. This will allow us to make good use of the Riesz–Thorin interpolation theorem that we discussed in Section 2.4.

Prelude: Next we present the famous result of Marcel Riesz from 1926. People had been struggling for years to prove that the Hilbert transform was bounded on the L^p spaces other than $p = 2$, so Riesz's result must be considered a true breakthrough. The actual argument that we now present is due to Salomon Bochner. But Riesz had slightly different tricks that also yielded a boundedness result just for the even, integer values of p. It requires an extra idea, namely interpolation of linear operators, to get the result for all p, $1 < p < \infty$ (as in the ensuing theorem).

Proposition 2.5.1 *The Hilbert transform is bounded on $L^p(\mathbb{T})$ when $p = 2k$ is a positive, even integer.*

Theorem 2.5.2 *The Hilbert transform is bounded on L^p, $1 < p < \infty$.*

Remark: The argument at the end of the proof of the last theorem (see Appendix 1) is commonly called a "duality argument." Later in the book, when this idea is needed, it will be invoked without further comment or detail.

We complete our consideration of the Hilbert transform by treating what happens on the spaces L^1 and L^∞.

Prelude: The failure of singular integrals on the extreme spaces L^1 and L^∞ is a fundamental part of the theory. The former fact gave rise, in part, to the relatively new idea of real-variable H^1_{Re} (the real-variable Hardy space—see Section 8.8). Singular integrals *are* bounded on H^1_{Re}. The latter fact gave rise to the space *BMO* of functions of bounded mean oscillation (also see Chapter 8). Singular integrals are also bounded on *BMO*. The book [KRA5] gives a sketch of some of these ideas. Stein's early work [STE5] on the space $L \log L$ (the space of functions f such that $\int |f| \log^+ |f| dx$ is finite) was another attempt to deal with the failure of singular integrals on L^1.

Proposition 2.5.3 *Norm summability for Fourier series fails in both L^1 and L^∞.*

The proof of this last fact is just another instance of the concept of duality, as noted earlier.

We conclude this discussion by noting that the Hilbert transform of the characteristic function of the interval [0, 1] is a logarithm function—do the easy calculation yourself. Thus the Hilbert transform does *not* map L^∞ to L^∞. By duality, it does not map L^1 to L^1. That completes our treatment of the nonboundedness of the Hilbert transform on these endpoint spaces.

2.6 The Modified Hilbert Transform

Capsule: The Hilbert transform, in its raw form, is a convolution operator with kernel $\cot \frac{t}{2}$. This is an awkward kernel to handle, just because it is a transcendental function. We show in this section that the kernel may be replaced by $1/t$. Most any question about the operator given by convolution with $\cot \frac{t}{2}$ may be studied by instead considering the operator given by convolution with $1/t$. Thus the latter operator has come (also) to be known as the Hilbert transform.

We repeat here a basic lesson from this chapter. We note that in practice, people do not actually look at the operator consisting of convolution with $\cot \frac{t}{2}$. This kernel is a transcendental function, and is tedious to handle. Thus what we do instead is to look at the operator

$$\widetilde{H} : f \longmapsto \text{P.V.} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-t) \cdot \frac{2}{t}\, dt. \tag{2.6.1}$$

Clearly the kernel $2/t$ is much easier to think about than $\cot \frac{t}{2}$. It is also homogeneous of degree -1, a fact that will prove significant when we adopt a broader point of view later.

Prelude: In the literature, people discuss variants of the Hilbert transform and still call these the "Hilbert transform." Once one understands the basic idea, it is a trivial matter to pass back and forth among all the different realizations of this fundamental singular integral.

Lemma 2.6.2 *If the modified Hilbert transform $\mathcal{H}$ is bounded on L^1, then it is bounded on L^∞.*

We end this section by recording what is perhaps the deepest result of basic Fourier analysis. Formerly known as the Lusin conjecture, and now as Carleson's theorem, this result addresses the pointwise convergence question for L^2. We stress that the approach to proving something like this is to study the *maximal Hilbert transform*—see Functional Analysis Principle I in Appendix 1.

Prelude: The next theorem is the culmination of more than fifty years of effort by the best mathematical analysts. This was *the* central question of Fourier analysis. Carleson's proof of the theorem was a triumph. Subsequently Fefferman [FEF4] produced another, quite different proof that was inspired by Stein's celebrated limits of sequences of operators theorem [STE6]. And there is now a third approach by Lacey and Thiele [LAT]. It must be noted that this last approach derives from ideas in Fefferman's proof.

Theorem 2.6.3 (Carleson [CAR]) *Let $f \in L^2(\mathbb{T})$. Then the Fourier series of f converges almost everywhere to f.*

The next result is based on Carleson's theorem, but requires significant new ideas.

Prelude: It definitely required a new idea for Richard Hunt to extend Carleson's result from L^2 to L^p for $1 < p < 2$ (of course the case L^p for $2 < p < \infty$ comes for free since then $L^p \subseteq L^2$). P. Sjölin [SJO1] has refined Hunt's theorem even further to obtain spaces of functions that are smaller than L^1, yet larger than L^p for every $p > 1$, on which pointwise convergence of Fourier series holds. The sharpest result along these lines is due to Hunt and Taibleson [HUT].

Theorem 2.6.4 (Hunt [HUN]) *Let $f \in L^p(\mathbb{T})$, $1 < p \leq \infty$. Then the Fourier series of f converges almost everywhere to f.*

A classical example of A. Kolmogorov (see [KAT], [ZYG]) provides an L^1 function whose Fourier series converges[4] *at no point* of $\mathbb{T}$. This phenomenon provides significant information: If instead the example were of a function with Fourier series diverging a.e., then we might conclude that we were simply using the wrong measure to detect convergence. But since there is an L^1 function with *everywhere diverging Fourier series*, we conclude that there is no hope for pointwise convergence in L^1.

Proofs of the Results in Chapter 2

Proof of Lemma 2.1.3: A rigorous proof of this lemma would involve a digression into distribution theory and the Schwartz kernel theorem. We refer the interested reader to either [STG1] or [SCH]. □

[4] It may be noted that Kolmogorov's original construction was very difficult. Nowadays, using functional analysis, this result may be had with little difficulty—see [KAT].

Proof of Proposition 2.1.4: Let f be a continuous real function on $[0, 2\pi)$. We normalize f (by subtracting off a constant) so that $\int f dx = 0$. Let u be its Poisson integral, so u is harmonic on the disk D and vanishes at 0. Let v be that harmonic conjugate of u on D such that $v(0) = 0$. Then $h = u + iv$ is holomorphic and $h(0) = 0$.

Fix $0 < r < 1$. Now we write

$$\begin{aligned}
0 &= 2\pi h^p(0) \\
&= \int_0^{2\pi} h^{2k}(re^{i\theta})\, d\theta \\
&= \int_0^{2\pi} [u(re^{i\theta}) + iv(re^{i\theta})]^{2k}\, d\theta \\
&= \int_0^{2\pi} u^{2k}\, d\theta + i\binom{2k}{1}\int u^{2k-1}v\, d\theta - \binom{2k}{2}\int u^{2k-2}v^2\, d\theta + \cdots \\
&\quad + i^{2k-1}\binom{2k}{2k-1}\int uv^{2k-1}\, d\theta + i^{2k}\int v^{2k}\, d\theta.
\end{aligned}$$

We rearrange the last equality as

$$\begin{aligned}
\int_0^{2\pi} v^{2k}\, d\theta \le \binom{2k}{2k-1}\int_0^{2\pi} |uv^{2k-1}|\, d\theta \\
+ \binom{2k}{2k-2}\int_0^{2\pi} |u^2v^{2k-2}|\, d\theta + \cdots \\
+ \binom{2k}{2}\int_0^{2\pi} |u^{2k-2}v^2|\, d\theta + \binom{2k}{1}\int_0^{2\pi} |u^{2k-1}v|\, d\theta \\
+ \int_0^{2\pi} |u^{2k}|\, d\theta.
\end{aligned}$$

We apply Hölder's inequality to each composite term on the right—using the exponents $2k/j$ and $2k/[2k - j]$ on the jth term, for $j = 1, 2, \ldots, 2k - 1$. It is convenient to let $S = [\int u^{2k}\, d\theta]^{1/2k}$ and $T = [\int v^{2k}\, d\theta]^{1/2k}$, and we do so. The result is

$$\begin{aligned}
T^{2k} \le \binom{2k}{2k-1} ST^{2k-1} + \binom{2k}{2k-2} S^2T^{2k-2} + \cdots \\
+ \binom{2k}{2} S^{2k-2}T^2 + \binom{2k}{1} S^{2k-1}T + S^{2k}.
\end{aligned}$$

Now define $U = T/S$ and rewrite the inequality as

$$U^{2k} \le \binom{2k}{2k-1} U^{2k-1} + \binom{2k}{2k-2} U^{2k-2} + \cdots + \binom{2k}{2} U^2 + \binom{2k}{1} U + 1.$$

Divide through by U^{2k-1} to obtain

$$U \le \binom{2k}{2k-1} + \binom{2k}{2k-2} U^{-1} + \cdots + \binom{2k}{2} U^{-2k+3} + \binom{2k}{1} U^{-2k+2} + U^{-2k+1}.$$

If $U \ge 1$, then it follows that

$$U \le \binom{2k}{2k-1} + \binom{2k}{2k-2} + \cdots + \binom{2k}{2} + \binom{2k}{1} + 1 \le 2^{2k}.$$

We conclude, therefore, that

$$\|v\|_{L^{2k}} \le 2^k \|u\|_{L^{2k}}.$$

But of course the function $v(re^{i\theta})$ is the Hilbert transform of $u(re^{i\theta})$. The proof is therefore complete. □

Proof of Theorem 2.1.5: We know that the Hilbert transform is bounded on L^2, L^4, L^6, ... We may immediately apply the Riesz–Thorin theorem (Section 2.1.3) to conclude that the Hilbert transform is bounded on L^p for $2 \le p \le 4$, $4 \le p \le 6$, $6 \le p \le 8$, etc. In other words, the Hilbert transform is bounded on L^p for $2 \le p < \infty$.

Now let $f \in L^p$ for $1 < p < 2$. Let φ be any element of $L^{p/[p-1]}$ with norm 1. Notice that $2 < p/[p-1] < \infty$. Then

$$\begin{aligned}
\int Hf \cdot \varphi d\theta &= \int \left[\int f(\psi) \cot \frac{\theta - \psi}{2} \, d\psi \right] \varphi(\theta) \, d\theta \\
&= \iint \varphi(\theta) \cot \frac{\theta - \psi}{2} \, d\theta f(\psi) \, d\psi \\
&= - \int \left[\int \varphi(\theta) \cot \frac{\psi - \theta}{2} \, d\theta \right] f(\psi) \, d\psi \\
&= - \int H\varphi(\psi) f(\psi) \, d\psi.
\end{aligned}$$

Using Hölder's inequality together with the fact that we know that the Hilbert transform is bounded on $L^{p/[p-1]}$, we may bound the right-hand side by the expression $C\|\varphi\|_{L^{p/[p-1]}} \|f\|_{L^p} \le \|f\|_{L^p}$. Since this estimate holds for any such choice of φ, the result follows. □

Proof of Proposition 2.1.6: It suffices for us to show that the modified Hilbert transform (as defined in Section 10.2) fails to be bounded on L^1 and fails to be bounded on L^∞. In fact, the following lemma will cut the job by half:

Proof of Lemma 2.2.1: Let f be an L^∞ function. Then

$$\|\mathcal{H}f\|_{L^\infty} = \sup_{\substack{\phi \in L^1 \\ \|\phi\|_{L^1} = 1}} \left| \int \mathcal{H}f(x) \cdot \phi(x) \, dx \right| = \sup_{\substack{\phi \in L^1 \\ \|\phi\|_{L^1} = 1}} \left| \int f(x) (\mathcal{H}^* \phi)(x) \, dx \right|.$$

But an easy formal argument (as in the proof of Theorem 10.5) shows that

$$\mathcal{H}^*\phi = -\mathcal{H}\phi.$$

Here $\mathcal{H}^*$ is the adjoint of $\mathcal{H}$. [In fact, a similar formula holds for *any* convolution operator—exercise.] Thus the last line gives

$$\begin{aligned}
\|\mathcal{H}f\|_{L^\infty} &= \sup_{\substack{\phi\in L^1 \\ \|\phi\|_{L^1}=1}} \left|\int f(x)\mathcal{H}\phi(x)\,dx\right| \\
&\leq \sup_{\substack{\phi\in L^1 \\ \|\phi\|_{L^1}=1}} \|f\|_{L^\infty}\cdot\|\mathcal{H}\phi\|_{L^1} \\
&\leq \sup_{\substack{\phi\in L^1 \\ \|\phi\|_{L^1}=1}} \|f\|_{L^\infty}\cdot C\|\phi\|_{L^1} \\
&= C\cdot\|f\|_{L^\infty}.
\end{aligned}$$

Here C is the norm of the modified Hilbert transform acting on L^1. We have shown that if $\mathcal{H}$ is bounded on L^1, then it is bounded on L^∞. That completes the proof. $\square$

Springer

http://www.springer.com/978-0-8176-4668-4

Explorations in Harmonic Analysis
With Applications to Complex Function Theory and the Heisenberg Group
Krantz, S.G.
2009, XIV, 362 p., Hardcover
ISBN: 978-0-8176-4668-4
A product of Birkhäuser Basel